

Robust Least Mean Mixed Norm Adaptive Filtering for α -Stable Random Processes

Gül Aydın¹, Oğuz Tanrıku², A. Enis Çetin¹

¹Bilkent University,

Department of Electrical Engineering,

06533, Bilkent, Ankara, Turkey

²Imperial College of Science, Technology and Medicine,

Department of Electrical and Electronic Engineering

Exhibition Road

London SW7 2BT, United Kingdom

ABSTRACT

Based on the concept of Fractional Lower Order Statistics (FLOS), we present the Robust Least Mean Mixed Norm (RLMMN) adaptive algorithm for applications in impulsive environments modeled by α -stable distributions. A sufficient condition for finite variance of the update term is obtained for the underlying α -stable process. Simulation results are provided regarding the identification of the parameters of an AR system.

I. INTRODUCTION

In many signal processing applications the noise is modeled as a Gaussian stochastic process. This assumption has been broadly accepted due to the Central Limit Theorem but a large class of physical observations such as low frequency atmospheric noise, man-made noise and underwater acoustic noise exhibit non-Gaussian behaviour [1]-[4]. A class of statistical models applicable in these cases is the α -stable ($0 < \alpha \leq 2$) distributions [5]. Having heavier tails than the Gaussian distributions, these distributions possess sharp spikes or occasional bursts in their realizations.

The α -stable distributions do not have closed form probability density functions except the cases $\alpha = 1$ (Cauchy distribution) and $\alpha = 2$ (Gaussian distribution). However they have closed form characteristic functions given by :

$$\phi(t) = \exp\{ist - \gamma|t|^\alpha[1 + i\beta\text{sgn}(t)w(t, \alpha)]\} \quad (1)$$

where α ($0 < \alpha \leq 2$) is the characteristic exponent, s is the location parameter, β ($-1 \leq \beta \leq 1$) is the index of skewness, $\gamma > 0$ is the dispersion parameter, $\text{sgn}(\cdot)$ denotes the signum function, and

$$w(t, \alpha) = \begin{cases} \tan\left(\frac{\alpha\pi}{2}\right) & \text{if } \alpha \neq 1 \\ \frac{2}{\pi} \log|t| & \text{if } \alpha = 1. \end{cases} \quad (2)$$

The distribution is called Symmetric α -Stable (SaS), if $\beta = 0$. The parameter α controls the tails of the

distribution. For $0 < \alpha < 2$ the distributions have algebraic tail which are significantly heavier than the exponential tail of the Gaussian distribution. The smaller the value of the α , the heavier the tails of the distributions. This property makes the α -stable distributions an appealing model for impulsive noise environments. It is well known that if the noise is impulsive, algorithms developed under the Gaussianity assumption may produce unacceptable results, [5], [6].

The Gaussian signals can be treated in a Hilbert space framework which allows the use of L_2 norm in various optimization criteria. Whereas the linear vector space generated by α -stable distributions is a Banach space when $1 \leq \alpha < 2$. In the Banach space of α -stable processes only L_p norm exists for $p < \alpha$, hence the use of L_2 norm is unacceptable in this case, [5], [7].

Due to the heavy tails, stable distributions do not have finite second or higher order moments, except the limiting case of $\alpha = 2$. Let X be an α -stable random variable with $0 < \alpha < 2$, then

$$\mathbf{E}[|X|^p] = \infty, \quad \text{if } p \geq \alpha. \quad (3)$$

However, for $0 \leq p < \alpha$ the Fractional Lower Order Moment (FLOM) is finite, i.e.,

$$\mathbf{E}[|X|^p] < \infty, \quad \text{if } 0 \leq p < \alpha. \quad (4)$$

If $\alpha = 2$, then

$$\mathbf{E}[|X|^p] < \infty, \quad \text{for all } p \geq 0. \quad (5)$$

The fractional p^{th} order moment [5] of an SaS random variable with zero location parameter, $a = 0$, is given by :

$$\mathbf{E}[|X|^p] = C(p, \alpha)\gamma^{p/\alpha} \quad \text{for } 0 < p < \alpha, \quad (6)$$

where

$$C(p, \alpha) = \frac{2^{p+1}\Gamma(\frac{p+1}{2})\Gamma(-p/\alpha)}{\alpha\pi\Gamma(-p/2)}. \quad (7)$$

In Equation (7), $\Gamma(\cdot)$ denotes the gamma function. The constant $C(p, \alpha)$ depends only on α and p .

In this paper, based on the concept of FLOS, we introduce a new family of adaptive filtering algorithms that utilize a linear mixture of moments similar to the LMMN algorithm [8,9]. It is assumed that the α -stable noise has finite mean corresponding $1 \leq \alpha < 2$.

II. THE LMMN ALGORITHM IN IMPULSIVE ENVIRONMENTS

The LMMN algorithm [8] is proposed for the cases when the noise distribution is a mixture of short and long-tail distributions. This algorithm combines the Least Mean Square (LMS) [10] and Least Mean Fourth (LMF) [11] algorithms. The cost function and the update equation of the LMMN algorithm are respectively given by

$$J_k = \frac{1}{2} \lambda E[e_k^2] + \frac{1}{4} (1 - \lambda) E[e_k^4] \quad (8)$$

and

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu e_k (\lambda + (1 - \lambda) e_k^2) \mathbf{X}_k. \quad (9)$$

The tap weights of the adaptive filter at time k are $\mathbf{W}_k = [w_{0,k} \dots w_{M-1,k}]^T$. The input vector is $\mathbf{X}_k = [x_k \dots x_{k-M+1}]^T$. The error signal is $e_k = d_k - \mathbf{W}_k^T \mathbf{X}_k$ where d_k is the desired signal. The step-size μ which controls the speed of convergence and $0 \leq \lambda \leq 1$ is the mixing parameter.

In Appendix-I, the update equation of the LMMN algorithm is studied for α -stable distributions. It is shown that the variance of the gradient vector is unbounded, i.e.,

$$E[\|\mathbf{W}_{k+1} - \mathbf{W}_k\|_2^2] = \infty. \quad (10)$$

which corresponds to unstable behaviour that agrees precisely with our experimental evaluations.

II.1 Simulation Studies

Consider an $AR(M)$ α -stable process, defined as follows,

$$x_k = \sum_{i=1}^M b_i x_{k-i} + u_k \quad (11)$$

where u_k is a symmetric α -stable (S α S) distribution. If $\{b_i\}$ is an absolutely summable sequence, then x_k also has a symmetric S α S distribution with the same parameters as u_k [5, 12].

An $AR(2)$ process with parameters $b_1 = 0.99$ and $b_2 = -0.1$ is identified by using the LMMN algorithm, where input signal x_k is generated from a S α S

process with $\alpha = 1.2$. The transient behaviour of the adaptive weights are illustrated in Fig. 1 for one realization. The mixing parameter values are chosen as $\lambda = 0.1, 0.3, 0.6$, and 0.9 and the step-size is $\mu = 5.2 \times 10^{-5}$. The LMMN algorithm experiences divergence for all values of λ even if very small step-sizes are used.

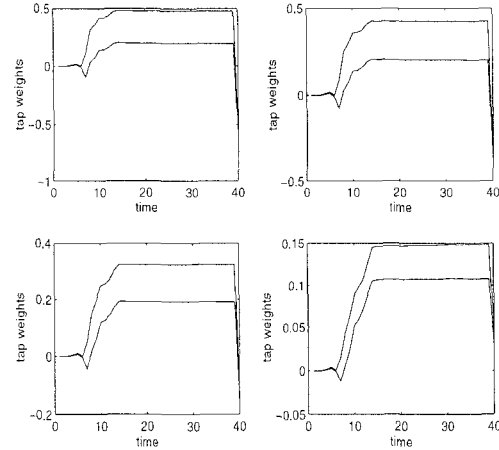


Fig. 1. Transient behaviour of tap weights for the LMMN algorithm for a S α S input with $\alpha = 1.2$. Divergence in all cases with $\lambda = 0.1, 0.3, 0.6, 0.9$.

III. ROBUST LMMN ADAPTIVE FILTERING

In order to ensure finite variance, i.e.,

$$E[\|\mathbf{W}_{k+1} - \mathbf{W}_k\|_2^2] < \infty \quad (12)$$

of the gradient vector for α -stable distributions, the LMMN algorithm is modified by using the concept of FLOS. In particular, the fact that in impulsive noise, the variance of the fractional lower order error

$$e_k^{} \triangleq |e_k|^b \text{sgn}(e_k) \text{ for } 0 < b < \alpha, \quad (13)$$

is bounded. Based on this observation we introduce the Robust-LMMN (RLMMN) algorithm as :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu e_k^{<a>} (\lambda + (1 - \lambda) \{e_k^{<a>}\}^2) \tilde{\mathbf{X}}_k \quad (14)$$

where $\tilde{\mathbf{X}}_k = [x_k^{<a>} \dots x_{k-M+1}^{<a>}]^T$. Unlike the LMMN algorithm, the RLMMN algorithm is expected to yield well-behaved convergence in impulsive noise environments. Note that, the actual weights define a stable stationary point of iterations for the RLMMN algorithm.

In Appendix-II, it is shown that a sufficient condition for the validity of the condition in (12) is

$$a < \alpha/8. \quad (15)$$

III.1 Simulation Studies

The FLOS based RLMMN algorithm is compared to the Normalized Least Mean-p Norm (NLMP) [7] algorithm. The update equation of the NLMP algorithm is given by :

$$\mathbf{W}_{k+1} = \mathbf{W}_k + \mu \frac{e_k^{<p-1>}}{\|\mathbf{X}_k\|_p^p + \varepsilon} \mathbf{X}_k \quad (16)$$

where $1 \leq p < \alpha$, $\varepsilon > 0$ and μ is chosen such that the NLMP algorithm is stable.

In Fig. 2, the results of a system identification experiment are presented. Note that $\alpha = 1.2$ and $\lambda = 0.9$. An AR(5) process with parameters $b_1 = 0.890$, $b_2 = -0.152$, $b_3 = 0.100$, $b_4 = -0.197$ and $b_5 = 0.097$ is generated which is driven by an i.i.d. SαS random process, u_k . The system mismatch is obtained by averaging $\|\mathbf{W}_k - \mathbf{W}_*\|_2^2$ over 100 Monte Carlo trials, where \mathbf{W}_k and \mathbf{W}_* are respectively the current tap weight vector and the optimal solution. The performance of the RLMMN algorithm is comparable with the NLMP algorithm depending on the mixture parameter. Note that, the step-sizes are chosen as $\mu = 5.2 \times 10^{-3}$ for both algorithms and the steady-state system mismatches are approximately equal.

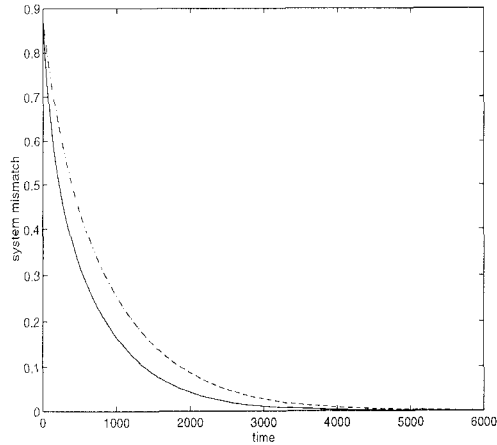


Fig. 2. The system mismatch for RLMMN algorithm (solid) and the NLMP algorithm (dashed) for $\alpha = 1.2$ and $\lambda = 0.9$.

IV. CONCLUSION

The Robust-LMMN adaptive algorithm is proposed for the applications where the signals have an impulsive nature. A bound is obtained on the variance of the gradient vector of the proposed algorithm. Simulation results are presented for adaptive system identification.

V. APPENDIX-I

The unbounded behaviour of the LMMN algorithm for α -stable processes is proven below. Let

$$e_k = d_k - \mathbf{W}_k^T \mathbf{X}_k = \mathbf{V}_k^T \mathbf{X}_k \quad (17)$$

where

$$\mathbf{V}_k \triangleq \mathbf{W}_* - \mathbf{W}_k \quad (18)$$

is the *weight error vector*. Since $\mathbf{V}_k - \mathbf{V}_{k+1} = \mathbf{W}_{k+1} - \mathbf{W}_k$, the weight update equation of the LMMN algorithm can be written in terms of \mathbf{V}_k as :

$$\mathbf{V}_{k+1} = \mathbf{V}_k - \mu e_k (\lambda + (1 - \lambda) e_k^2) \mathbf{X}_k \quad (19)$$

and

$$\mathbf{E}[\|\mathbf{W}_{k+1} - \mathbf{W}_k\|_2^2] =$$

$$\mu^2 \mathbf{E}[(\mathbf{V}_k^T \mathbf{X}_k)^2 (\lambda + (1 - \lambda) (\mathbf{V}_k^T \mathbf{X}_k)^2) \|\mathbf{X}_k\|_2^2]. \quad (20)$$

Having infinite value of (20) can be simply shown by considering the term, $\mathbf{E}[(\mathbf{V}_k^T \mathbf{X}_k)^2 \|\mathbf{X}_k\|_2^2]$, after expanding the parenthesis. By using the “independence theory assumption” in [10] and assuming that \mathbf{V}_k and \mathbf{X}_k are independent, we have

$$\begin{aligned} \mathbf{E}[(\mathbf{V}_k^T \mathbf{X}_k)^2 \|\mathbf{X}_k\|_2^2] &= \sum_{i=1}^M \mathbf{E}[v_{k-i}^2] \mathbf{E}[x_{k-i}^2 \|\mathbf{X}_k\|_2^2] \\ &+ \sum_{i=1}^M \sum_{j=1, j \neq i}^M \mathbf{E}[v_{k-i} v_{k-j}] \mathbf{E}[x_{k-i} x_{k-j} \|\mathbf{X}_k\|_2^2]. \end{aligned} \quad (21)$$

We have $\mathbf{E}[x_{k-i}^2 \|\mathbf{X}_k\|_2^2] = \infty$, since $\mathbf{E}[x_{k-i}^4] = \infty$ within this expectation and therefore, (10) is proven for LMMN algorithm under α -stable distributions.

VI. APPENDIX-II

In order to prove (15) the update equation of the RLMMN algorithm can be rewritten in terms of \mathbf{V}_k as :

$$\begin{aligned} \mathbf{V}_{k+1} &= \mathbf{V}_k - \mu (\mathbf{V}_k^T \mathbf{X}_k)^{<a>} \\ &\times \left[\lambda + (1 - \lambda) \{ (\mathbf{V}_k^T \mathbf{X}_k)^{<a>} \}^2 \right] \tilde{\mathbf{X}}_k. \end{aligned} \quad (22)$$

Therefore we have :

$$\|\mathbf{W}_{k+1} - \mathbf{W}_k\|_2^2 = \mu^2 \{ (\mathbf{V}_k^T \mathbf{X}_k)^{<a>} \}^2$$

$$\times \left[\lambda + (1 - \lambda) \{ (\mathbf{V}_k^T \mathbf{X}_k)^{<a>} \}^2 \right]^2 \|\tilde{\mathbf{X}}_k\|_2^2. \quad (23)$$

By using the Hölder inequality,

$$|\mathbf{V}_k^T \mathbf{X}_k| \leq \|\mathbf{V}_k\|_1 \|\mathbf{X}_k\|_\infty \quad (24)$$

in (23), we obtain

$$\begin{aligned} \|\mathbf{W}_{k+1} - \mathbf{W}_k\|_2^2 &\leq \mu^2 \|\mathbf{V}_k\|_1^{2a} \|\mathbf{X}_k\|_\infty^{2a} [\lambda^2 + 2\lambda(1 - \lambda)] \\ &\quad \|\mathbf{V}_k\|_1^{2a} \|\mathbf{X}_k\|_\infty^{2a} + (1 - \lambda)^2 \|\mathbf{V}_k\|_1^{4a} \|\mathbf{X}_k\|_\infty^{4a} \|\tilde{\mathbf{X}}_k\|_2^2 \end{aligned}$$

If the statistical independence assumption of \mathbf{V}_k and \mathbf{X}_k is used once more, we have

$$\begin{aligned} \mathbf{E}[\|\mathbf{W}_{k+1} - \mathbf{W}_k\|_2^2] &\leq \\ &\mu^2 [\lambda^2 \mathbf{E}[\|\mathbf{V}_k\|_1^{2a}] \mathbf{E}[\|\mathbf{X}_k\|_\infty^{2a} \|\tilde{\mathbf{X}}_k\|_2^2] \\ &\quad + 2\lambda(1 - \lambda) \mathbf{E}[\|\mathbf{V}_k\|_1^{4a} \|\mathbf{X}_k\|_\infty^{4a} \|\tilde{\mathbf{X}}_k\|_2^2] \\ &\quad + (1 - \lambda)^2 \mathbf{E}[\|\mathbf{V}_k\|_1^{6a}] \mathbf{E}[\|\mathbf{X}_k\|_\infty^{6a} \|\tilde{\mathbf{X}}_k\|_2^2]]. \end{aligned} \quad (25)$$

Since $x_{k-j} \triangleq \|\mathbf{X}_k\|_\infty < \infty$, from (25), we obtain

$$a < \alpha/8 \quad (26)$$

as a sufficient condition for the fulfillment of the condition in (12). It is assumed that no parameterize divergence occurs in the RLMMN algorithm, i.e., $\sup_k \|\mathbf{V}_k\|_\infty < \infty$. In this case, even if \mathbf{V}_k is a S α S vector process, when (26) is valid, we have $\mathbf{E}[\|\mathbf{V}_k\|_1^{2a}] < \infty$, $\mathbf{E}[\|\mathbf{V}_k\|_1^{4a}] < \infty$ and $\mathbf{E}[\|\mathbf{V}_k\|_1^{6a}] < \infty$.

VII. REFERENCES

- [1] B. Mandelbrot and J.W. Van Ness, "Fractional Brownian Motions, Fractional Noises, and Applications," *SIAM Review*, Vol. 10, pp. 422-437, 1968.
- [2] S.S. Pillai, M. Harisankar, "Simulated performance of a DS spread spectrum system in impulsive atmospheric noise," *IEEE Trans. Electromagnetic Compat.*, Vol. 29, pp. 80-82, 1987.
- [3] M. Bouvet and S.C. Schwartz, "Comparison of adaptive and robust receivers for signal detection in ambient underwater noise," *IEEE Trans. Acous. Speech and Signal Proc.*, Vol. 37, pp. 621-626, 1989.
- [4] D. Middleton, "Statistical Physical Models of Electromagnetic interference," *IEEE Transactions on Electromagnetic Compat.*, Vol. 19, pp. 106-127, 1977.
- [5] M. Shao and C.L. Nikias, "Signal Processing with Fractional Lower Order Moments : Stable Processes and their Applications," *Proc. IEEE*, Vol. 81, pp. 986-1009, 1993.
- [6] S. A. Kassam, *Signal Detection in Non-Gaussian Noise*. New York: Springer, 1988.
- [7] O. Arikan, A. E. Çetin, and E. Erzin, "Adaptive Filtering for Non-Gaussian Stable Processes," *IEEE Transaction on Signal Processing Letters*, Vol. 1, pp. 1-3, Nov. 1994.
- [8] O. Tanrikulu, J.A. Chambers and A.G. Constantinides, "Least Mean Mixed-Norm Adaptive Filtering," *Electronics Letters*, Vol. 30, No. 19, pp. 1574-1575, Sep. 1994.
- [9] O. Tanrikulu and J.A. Chambers, "Convergence and steady-state properties of the Least Mean Mixed Norm (LMMN) adaptive algorithm", *to appear in IEE Proceedings-VIS*, 1996.
- [10] S. Haykin *Adaptive Filter Theory*, Prentice Hall, NJ, 1986.
- [11] E. Walach, and B. Widrow : "The Least Mean Fourth (LMF) adaptive algorithm and its family," *IEEE Trans.* 1984, IT-30, pp. 275-283.
- [12] Y. Hosoya, "Discrete Time Stable Processes and Their Certain Properties," *Ann. Prob.*, Vol. 6, No. 1, pp. 94-105, 1978.